Title

Functional variants of DOG1 control seed chilling responses and variation in seasonal life-history strategies in Arabidopsis thaliana.

Journal

Proceedings of the National Academy of Sciences of the United States of America, 117(5)

ISSN

0027-8424

Authors

Martínez-Berdeja, Alejandra
Stitzer, Michelle C
Taylor, Mark A
et al.

Publication Date

2020-02-01

DOI

10.1073/pnas.1912451117

Peer reviewed

Functional variants of *DOG1* control seed chilling responses and variation in seasonal life-history strategies in *Arabidopsis thaliana*

Alejandra Martínez-Berdeja[a], Michelle C. Stitzer[a,b], Mark A. Taylor[a], Miki Okada[a], Exequiel Ezcurra[c], Daniel E. Runcie[d], and Johanna Schmitt[a,b,1]

[a]Department of Evolution and Ecology, University of California, Davis, CA 95616; [b]Center for Population Biology, University of California, Davis, CA 95616; [c]Department of Botany and Plant Sciences, University of California, Riverside, CA 92521; and [d]Department of Plant Sciences, University of California, Davis, CA 95616

Contributed by Johanna Schmitt, November 22, 2019 (sent for review July 19, 2019; reviewed by Leonie Bentsink, Xavier Picó, and Peter Tiffin)

The seasonal timing of seed germination determines a plant's realized environmental niche, and is important for adaptation to climate. The timing of seasonal germination depends on patterns of seed dormancy release or induction by cold and interacts with flowering-time variation to construct different seasonal life histories. To characterize the genetic basis and climatic associations of natural variation in seed chilling responses and associated life-history syndromes, we selected 559 fully sequenced accessions of the model annual species *Arabidopsis thaliana* from across a wide climate range and scored each for seed germination across a range of 13 cold stratification treatments, as well as the timing of flowering and senescence. Germination strategies varied continuously along 2 major axes: 1) Overall germination fraction and 2) induction vs. release of dormancy by cold. Natural variation in seed responses to chilling was correlated with flowering time and senescence to create a range of seasonal life-history syndromes. Genome-wide association identified several loci associated with natural variation in seed chilling responses, including a known functional polymorphism in the self-binding domain of the candidate gene *DOG1*. A phylogeny of *DOG1* haplotypes revealed ancient divergence of these functional variants associated with periods of Pleistocene climate change, and Gradient Forest analysis showed that allele turnover of candidate SNPs was significantly associated with climate gradients. These results provide evidence that *A. thaliana*'s germination niche and correlated life-history syndromes are shaped by past climate cycles, as well as local adaptation to contemporary climate.

delay of germination | germination niche | seed dormancy | genome-wide association | stratification

The seasonal timing of seed germination is a niche construction trait that shapes phenology and life history throughout the life cycle (1–10), and may be critical for adaptation to climate (11–13). Germination timing is under strong selection in nature (1, 5, 13–17) and determines the selective environment for subsequent life-history traits, such as flowering time (1, 15, 16, 18), driving correlated evolution of suites of life-history traits across climate gradients (11, 12, 19, 20). Seasonal germination timing in the wild depends upon natural variation in primary dormancy at dispersal and in responses to environmental cues that mediate annual cycles of dormancy release and secondary dormancy induction. In particular, variation in seed responses to seasonal chilling cues can determine whether seeds germinate in fall or spring (21–24). Fall vs. spring germination defines distinct life-history strategies whose adaptive value may vary across climates (11, 12). To understand the role of these critical niche construction traits in climate adaptation, it is important to describe their genetic basis. Here we characterize natural variation in seed chilling responses and correlated life-history traits in the annual species *Arabidopsis thaliana*, identify genomic variants underlying these traits, and test whether these variants exhibit a signature of adaptation to climate across the species range.

In annual plants, variation in seasonal germination timing creates alternative life-history strategies (*SI Appendix*, Fig. S1). Winter annuals germinate in autumn, overwinter, and then flower and disperse seed in spring, whereas spring or summer annuals overwinter as seeds and germinate, flower, and disperse seed in spring or summer (11, 17, 21, 23). Mixtures of fall and spring germination cohorts are also observed within populations (7, 10, 11, 17, 21), a form of within-year bet hedging (25). Whether and when seeds germinate in a given seasonal environment depends upon the germination niche, the range of conditions under which germination is possible (7, 11, 21, 26). Primary dormancy of freshly dispersed seeds varies among genotypes and seed maturation environments (22, 23, 27–33). Release from primary dormancy may occur through after-ripening at warm ambient temperatures, or through short exposure to chilling (21, 22, 34, 35). Both of these mechanisms allow germination of winter annual seedlings in fall. In many winter annuals, seeds remaining in the soil seed bank in fall enter secondary dormancy when exposed to prolonged winter

Significance

The seasonal timing of seed germination is critical for plant fitness in different climates. To germinate at the right time of year, seeds respond to seasonal environmental cues, such as cold temperatures. We characterized genetic variation in seed dormancy responses to cold across the geographic range of a widespread annual plant. Induction of secondary seed dormancy during winter conditions (which restricts germination to autumn) was positively correlated with flowering time, constructing winter and spring seasonal life-history strategies. Variation in seed chilling responses was strongly associated with functional variants of a known dormancy gene. These variants showed evidence of ancient diversification associated with Pleistocene glacial cycles, and were associated with climate gradients across the species' geographical range.

Author contributions: A.M.-B. and J.S. designed research; A.M.-B. and M.O. performed research; A.M.-B., M.C.S., M.A.T., E.E., D.E.R., and J.S. analyzed data; and A.M.-B., M.C.S., and J.S. wrote the paper.

Reviewers: L.B., Wageningen University & Research; X.P., Estación Biológica de Doñana (Consejo Superior de Investigaciones Cientifica); and P.T., University of Minnesota.

The authors declare no competing interest.

Data deposition: Germination, phenology and *DOG1* haplotype data are available on Dryad repository, https://doi.org/10.25338/B8VS4P. *DOG1* haplotype data and code are also available on GitHub, https://github.com/mcstitzer/martinez-berdeja_dog1.

See online for related content such as Commentaries.

[1]To whom correspondence may be addressed. Email: jschmitt@ucdavis.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1912451117/-/DCSupplemental.

First published January 21, 2020.

chilling, preventing winter and spring germination (21, 22, 26). In contrast, summer annual seeds lose dormancy with chilling over winter, allowing germination in spring (21, 23). Seeds that lose dormancy in fall and do not enter secondary dormancy over the winter may germinate in either fall or spring resulting in a mixture of fall and spring annual life histories. Natural variation in dormancy responses to seasonal chilling cues thus determines the germination niche and the expression of seasonal life-history strategies within and among populations in different environments.

Germination timing also determines the selective environment for subsequent life-history traits, such as flowering time (1, 15, 16, 18). Seedlings that germinate in fall may be selected to delay flowering until winter is past, which may favor late flowering and strong vernalization (winter chilling) requirements. However, strong vernalization requirements and late flowering may be maladaptive for summer annuals, especially if the growing season is short. Consequently, natural selection on germination and later life-history traits may be correlated, and adaptation to climate across a species' range may therefore involve coordinated evolution of suites of life-history traits (11, 12, 19, 20, 36, 37).

To understand the adaptive evolution of the germination niche across different climates, it is important to elucidate its genetic basis. It is also of interest to examine the genetic basis of correlated traits that may form adaptive life-history syndromes, such as winter vs. summer annual life histories. If different traits share common genetic mechanisms, this pleiotropy may constrain or facilitate adaptive evolution of multivariate syndromes (38). On the other hand, if the loci contributing to variation in different traits do not overlap, then adaptive divergence of life-history syndromes could only occur through correlated selection response at different sets of loci across a climate gradient. Once loci underlying life-history variation are identified, we can test for a geographic signature of selection by climate. If allelic variants at those loci are involved in adaptation to climate, we expect them to be significantly associated with relevant climate variables across the landscape (39–41).

The model plant *A. thaliana* is an ideal system for dissecting the genetic and environmental determinants of climate adaptation in the germination niche and correlated life-history traits. Extensive sequence data are available for this annual species, facilitating genome-wide association (GWA) (42, 43). *A. thaliana* inhabits a wide native climate range across Eurasia and Africa, and has shifted across the landscape with Pleistocene climate cycles, repeatedly contracting and expanding out of glacial refugia (44–47). Across this range, the species exhibits substantial life-history variation (11, 13, 48). The pathways involved in seed dormancy and flowering time have been characterized through functional studies (24, 49). Natural variation in primary dormancy and after-ripening is also well documented (11, 12, 19, 20, 50), and underlying allelic variants have been identified (13, 27, 50–54). However, less is known about the genetic basis of natural variation in seed chilling responses in this species, despite their importance for fine-tuning dormancy cycles to shape seasonal germination timing (19, 22).

DELAY OF GERMINATION 1 (*DOG1*), a key regulator of seed dormancy and a member of a small gene family of unknown molecular function (27), is a particularly important candidate gene for natural variation in germination niche traits. This gene controls primary seed dormancy through multiple mechanisms (55–62). Allelic variants of *DOG1* are associated with natural variation in primary dormancy, after-ripening, and germination timing in the field (5, 13, 50, 52, 63). *DOG1* expression is associated with dormancy variation measured as after-ripening time (27, 53) and exhibits clinal variation; southern ecotypes have higher *DOG1* expression associated with longer after-ripening (12, 51, 63). *DOG1* variation is also associated with natural variation in flowering time (42, 43, 64) and may have pleiotropic effects both indirectly through cascading effects of germination timing (6) and directly through its effects on levels of associated micro RNAs (60). However, little is known about the contribution of *DOG1* functional polymorphisms to natural variation in germination responses to chilling.

Here we combine experimental phenotypic data from a geographically diverse set of accessions, whole-genome polymorphism data (43), and geographic and climate information to address the following questions:

1) What is the range of natural variation of germination responses to chilling? How does natural variation in germination responses to chilling covary with other life-history traits to shape winter and spring annual life-history syndromes across the landscape?
2) What is the genetic basis of natural variation in germination responses to chilling and associated life-history traits? Does allelic variation in *DOG1* contribute to variation in seed chilling responses to shape the seasonal germination niche?
3) Do these genetic variants exhibit a geographic signature of adaptation to climate?

Results and Discussion

Natural Variation in Seed Chilling Responses Creates Diverse Seasonal Life Histories. We grew 559 *Arabidopsis* accessions to reproductive maturity at 14 °C with a 12/12-h photoperiod following 6 wk of vernalization at 4 °C. Fresh seeds were harvested from each plant when siliques matured, and were tested for germination at 22 °C immediately (i.e., base germination), or after 4, 8, 11, 15, 22, or 32 d of dark stratification at 4 °C and 10 °C. We performed a principal component analysis (PCA) on all germination phenotypes. The first 2 principal components (PCs) of germination timing and dormancy level explained 91% of the germination variation ($PC1_{germ}$ = 83%, $PC2_{germ}$ = 7%) (Fig. 1 and *SI Appendix*, Fig. S2). We also measured days to flowering (DTF), and days to senescence (DTS) of the maternal plants. Additionally, we assayed germination of dry-stored seeds at 6-wk intervals to quantify after-ripening requirements (days of seed dry storage required to reach 50% germination, $DSDS_{50}$) (Table 1).

All germination variables measured were positively associated to $PC1_{germ}$, (*SI Appendix*, Table S1), and ecotypes with a positive score along this axis had overall high germination. Accessions with high overall germination ($PC1_{germ}$) were late-flowering and late-senescing and had short after-ripening times, whereas accessions with low germination were early-flowering and early-senescing and had long after-ripening times (Table 1). Variation in overall germination fraction reflects variation in primary dormancy, consistent with previous observations of natural variation in dormancy measured as after-ripening requirement (11, 12, 50). Accessions with high germination were distributed in central and northern Europe (*SI Appendix*, Fig. S3*A*). Low dormancy allows immediate germination after dispersal, which is favored by selection in northern climates (13) and may result in multiple generations per year in midlatitudes with wet summers (7, 65). In contrast, primary dormancy that cannot be broken by chilling exposure, correlated with strong after-ripening requirements, would maintain a large population in the seed bank (7), a potential bet-hedging strategy (66). Accessions with low germination were distributed in south and central Spain and southern Europe (*SI Appendix*, Fig. S3*A*).

$PC2_{germ}$ represents the germination response to cold. Base germination and germination percent of all stratification times at 10 °C were positively associated with $PC2_{germ}$, while germination percent at 11 to 32 d at 4 °C were negatively associated to this axis (*SI Appendix*, Table S1). Accessions with a positive score on $PC2_{germ}$ had high base germination and secondary dormancy induced by prolonged cold exposure at 4 °C (Fig. 1 and *SI Appendix*, Fig. S2 and Table S1), flowered and senesced later, and had longer after-ripening periods (Fig. 2 and Table 1). These accessions would behave as winter annuals, with germination

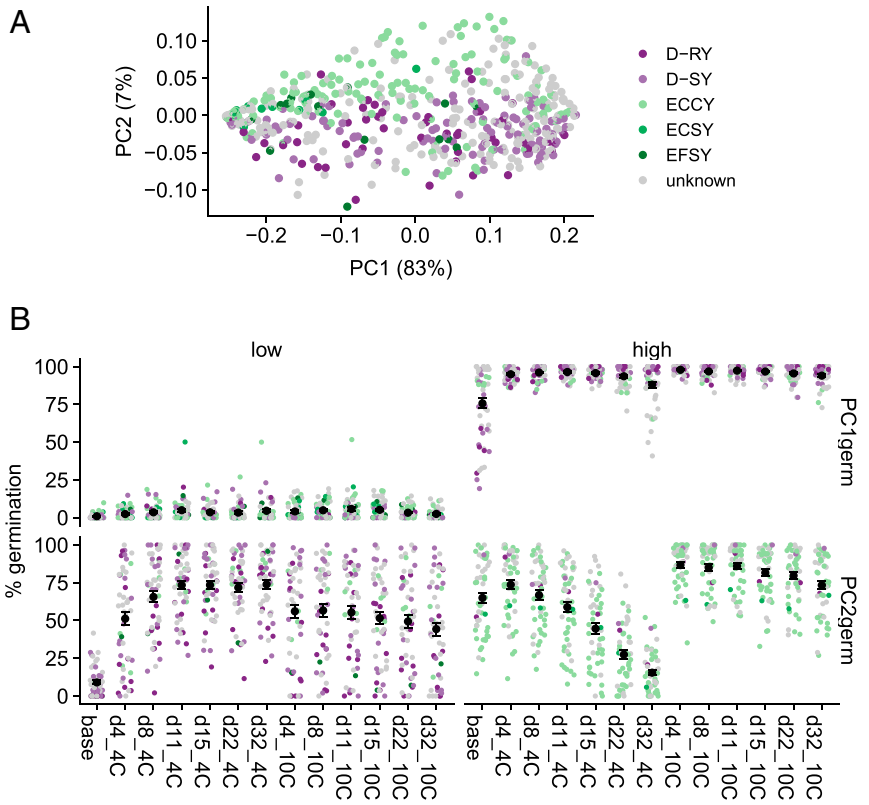

Fig. 1. (*A*) PC scores of germination variables (base, and cold treatments: 4, 8, 11, 15, 22, and 32 d at 4 and 10 °C). Percent variance explained by each axis: $PC1_{germ} = 83\%$, $PC2_{germ} = 7\%$. (*B*) Mean percent germination in each treatment for the lines with the 10% highest and lowest scores on $PC1_{germ}$ and $PC2_{germ}$. Colors indicate the amino acid sequence at the *DOG1* self-binding domain (53), which we use to define haplotypes, and grey color indicates accessions for which we were not able to assign a haplotype.

restricted to late summer or fall, and flowering and seed dispersal the following summer. Many of these accessions occur in Scandinavia and the Iberian Peninsula (*SI Appendix*, Fig. S3*B*). In contrast, accessions with a negative score on $PC2_{germ}$ that break primary dormancy with brief chilling exposure and do not enter secondary dormancy with prolonged chilling also displayed fast phenological transitions and had long after-ripening times, preventing summer germination. Dormancy release by cold would allow germination in both fall and spring, and coupled with rapid reproduction, would facilitate rapid cycling or spring annual life histories. Seeds that do not germinate in fall would overwinter in the soil under cold temperatures and would be ready to germinate the next spring (26). This seasonally flexible germination may be favored in disturbed, ruderal landscapes. These accessions were common at midlatitudes in Europe and England (*SI Appendix*, Fig. S3*B*).

The timing of germination defines the temporal environment an annual plant experiences, and interacts with later life-cycle traits to generate different life histories. Our results are consistent with previous observations of covariation in seed dormancy (measured as after-ripening requirements) and flowering-time traits across climatic gradients in *A. thaliana* (12, 19, 20, 37). Trait syndromes allow local adaptation, as covariation in life-history traits, growth rate, and stress responses influence fitness (37). However, our observations of germination responses to chilling add a new dimension to this picture of multivariate life histories, showing that cold-induced secondary dormancy is correlated with flowering and senescence to construct a wide range of seasonal life-history strategies.

Table 1. Correlations between the first 2 PCs from the germination variables PCA with phenology variables of the mother plants

		$PC1_{germ}$		$PC2_{germ}$	
Variable	df	*r*	*P* value	*r*	*P* value
DTF	523	−0.07	0.09	0.43	$<2.2 \times 10^{-16}$
DTS	523	0.18	5.19×10^{-5}	0.50	$<2.2 \times 10^{-16}$
$DSDS_{50}$	517	−0.49	$<2.2 \times 10^{-16}$	−0.43	$<2.2 \times 10^{-16}$

Genetic Architecture of Seed Responses to Chilling: A Major Role for DOG1. To understand the genetic architecture of the germination niche, we performed GWA analyses on the germination PCs using 498 fully sequenced *A. thaliana* accessions and 3,483,598 single nucleotide polymorphisms (SNPs; 1001 Genomes Consortium). A region on chromosome (Chr.) 5 was the most highly associated to both the first and second PCs of germination. This region included 1 SNP significantly associated with $PC1_{germ}$ and 42 SNPs significantly associated with $PC2_{germ}$. Several of these SNPs likely tag the same functional variant, as they are organized into only 4 linkage disequilibrium (LD) blocks spanning multiple candidate genes. Different SNPs were most highly associated to $PC1_{germ}$ and $PC2_{germ}$, although the closest gene to both was *DOG1* (Fig. 3 and *SI Appendix*, Tables S2 and S3). The SNP significantly associated to $PC1_{germ}$ and $PC2_{germ}$ (Chr. 5: 18,592,365) (*SI Appendix*, Table S2) had previously been shown to tag the promoter region of *DOG1* and to be associated with after-ripening time variation (50) (*SI Appendix*, Fig. S4). However, in our experiments, $DSDS_{50}$ was associated with a different set of SNPs in several LD blocks, which did not include *DOG1* (*SI Appendix*, Fig. S5 and Table S2). Using more permissive thresholds, we found that the top 1,000 SNPs

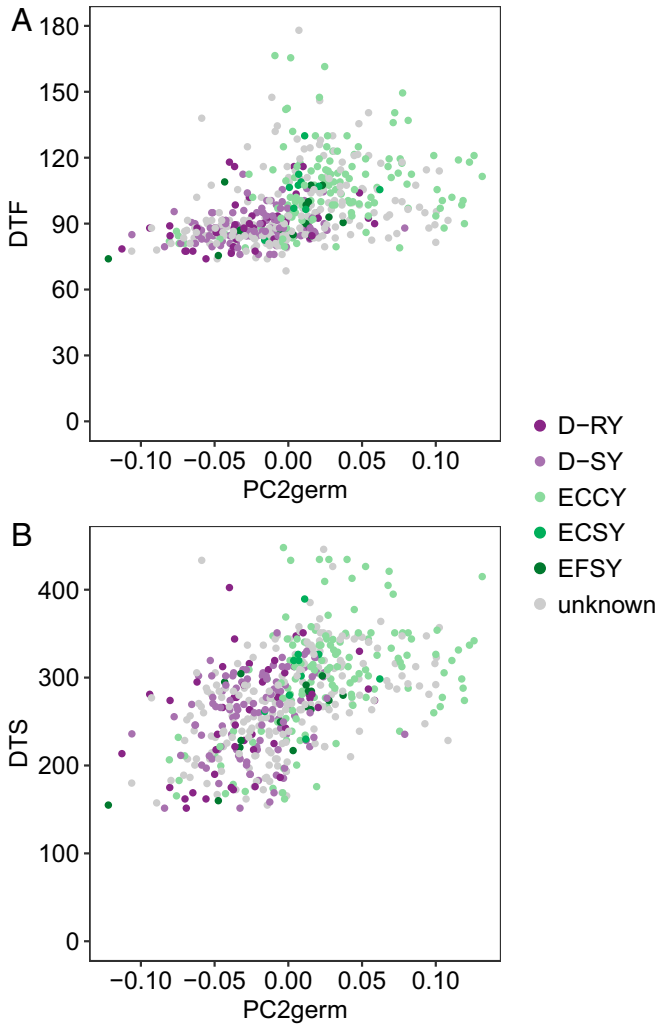

Fig. 2. Correlations between $PC2_{germ}$ and DTF (*A*) and DTS (*B*). Colors indicate the amino acid sequence at the self-binding domain for different *DOG1* haplotypes.

associated to each trait were enriched for seed-specific gene ontologies (*SI Appendix*, Tables S4 and S5).

As SNPs near *DOG1* were associated in both analyses, we further characterized genetic variation at *DOG1*. Known indel polymorphism within the self-binding domain of DOG1 (amino acid positions 13 to 16) (53) is not fully captured by published SNP data for *A. thaliana* (43, 45, 46). We therefore locally reassembled exon 1 of *DOG1* using raw sequence data from the 1001 Genomes Consortium (43), as well as African (45) and Chinese (46) *A. thaliana* accessions. This indel appears to be functional, as it likely affects dimerization of DOG1 proteins and alters their incorporation into larger protein complexes (53). Haplotypes in the self-binding domain have been defined by the amino acid variants present. We identified the ancestral self-binding *DOG1* E-haplotype (ECCY) and 2 loss-of-function *DOG1* D-haplotypes (D-RY, D-SY) previously reported (53), as well as 2 additional rare derived haplotypes without an amino acid deletion in the self-binding domain (ECSY and EFSY) found in less than 5% of individuals and restricted to Iberia (*SI Appendix*, Figs. S6 and S7). *DOG1* haplotype identities are significantly associated with $PC2_{germ}$ scores (*SI Appendix*, Table S6*A*). The loss-of-function D-RY and D-SY *DOG1* haplotypes share the reference allele of 3 of the most highly associated SNPs for $PC2_{germ}$ in the same LD block (Chr. 5: 18,590,327, 18,590,741, 18,590,743), while the alternative allele is associated to the self-binding E-haplotypes (*SI Appendix*, Fig. S8 and Table S2). Although these SNPs tag the indel and nonsynonymous variant that differentiates D- and E-containing haplotypes, our haplotype genotyping distinguishes further fine-scale genetic differences between accessions. We expect this is particularly important in refining the self-binding ability of the DOG1 protein, which is known to play an important role in the germination response to chilling (53, 59). Accessions carrying the self-binding *DOG1* E-haplotypes were characterized by strong secondary dormancy induced by prolonged chilling (high $PC2_{germ}$) (Fig. 2); in contrast, the loss-of-function *DOG1* D-haplotypes, had strong primary dormancy release and no secondary dormancy induced by cold (low $PC2_{germ}$) (Fig. 2).

We tested for epistatic effects of *DOG1* on germination PCs by dividing the data into 2 subsets of *DOG1* haplotypes (D-haplotypes, $n = 158$, and E-haplotypes, $n = 161$) and performing a GWA analysis within each haplotype. We found a peak on Chr. 5 for $PC2_{germ}$ for the E-haplotype annotated to AT5G65100, an ethylene-insensitive 3 family protein (*SI Appendix*, Fig. S9*A* and Table S7). A SNP associated to $PC2_{germ}$ in accessions carrying the E-haplotype falls within a region on Chr. 5 linked to a fitness quantitative trait loci related to seed coat alterations (67) and seed coat mucilage production (68). Despite not being significantly associated to $PC2_{germ}$ in the D-haplotype accessions, this SNP has a similar allele frequency among the D- and E-haplotypes (*SI Appendix*, Table S8). Thus, this association arises not because the causative variant is absent in 1 haplotypic background, but rather because different SNPs have different effect sizes in different backgrounds. The epistatic interaction between the *DOG1* E-haplotype and the associated allele in this region of Chr. 5 give additional insight into the complex gene interactions involved in the regulation of germination responses to chilling. Although not measured in our study, it is likely that *DOG1* expression levels are involved in the chilling response as well (12, 53, 57, 59, 62, 63, 69–73). Other mechanisms regulating *DOG1* expression might also be involved in the chilling germination responses: For example, the antisense transcript *asDOG1*, which negatively regulates the expression of *DOG1* (74, 75).

Life-History Syndromes Resulted from Correlated Selection on Multiple Loci. Accessions with loss-of-function *DOG1* D-haplotypes showed strong primary dormancy release and no secondary dormancy induced by cold, and also flowered and senesced early. In contrast, individuals with self-binding *DOG1* E-haplotypes showed secondary dormancy induced by prolonged cold, and flowered and senesced later on average (Fig. 2). Besides $PC2_{germ}$, *DOG1* haplotypes explained variation in DTF and DTS (*SI Appendix*, Fig. S10 and Table S6*A*). Similar D- and E-haplotype germination patterns were observed for geographically restricted samples within Spain and Sweden (*SI Appendix*, Figs. S11 and S12 and Tables S6 *B* and *C*). Pleiotropic effects of *DOG1* on flowering and germination mediated by miR156 and miR172 may be a mechanism behind these trait correlations (60, 76).

Our results also showed that GWAs for DTF were polygenic (39 SNPs organized in 11 LD blocks) (Fig. 3 and *SI Appendix*, Table S2) and SNPs tagging *DOG1* were not significantly associated with DTF. DTS showed no associated SNPs above a permutation threshold for significance (*SI Appendix*, Fig. S5 and Table S2). However, the 5 SNPs associated to $PC2_{germ}$ tagging *DOG1* have been previously associated to different flowering-time phenotypes under various conditions (Fig. 3 and *SI Appendix*, Fig. S13 and Table S2) (42, 43, 64, 77). The lack of association of DTF with *DOG1* in our study could partly be due to the accessions sampled, as we performed an additional GWA analysis using only our accessions with subsampled 1001 Genomes Consortium flowering phenotypes (43) and did not find evidence of association of DTF with SNPs tagging *DOG1* (*SI*

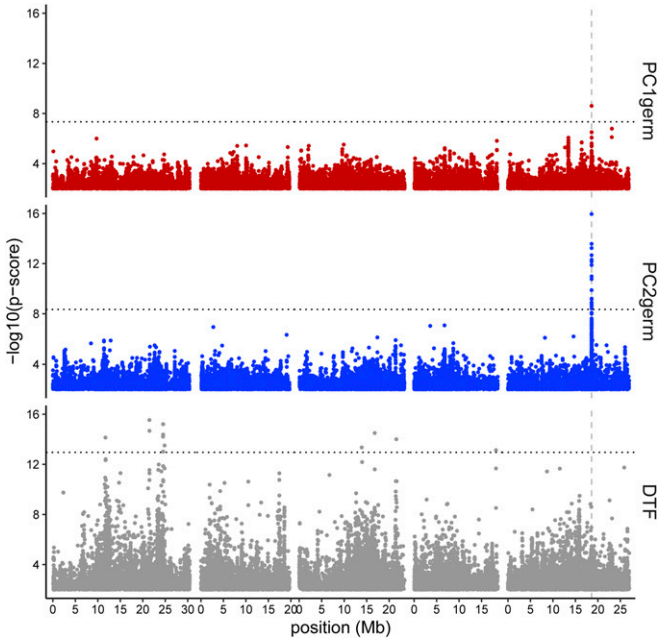

Fig. 3. Chromosome location (Mb) and association *P* value from GWAS for the SNPs on $PC1_{germ}$, $PC2_{germ}$, and DTF (permutation-based significance thresholds are shown by dotted horizontal lines). The vertical dashed grey line indicates the location of *DOG1* gene.

Appendix, Fig. S14). Our sampling includes mostly European accessions while the 1001 Genomes phenotypes also include Asian accessions (78). Moreover, the lack of association could also be due to the different vernalization and growth conditions used in each experiment.

We also tested for epistatic effects of *DOG1* haplotypes on flowering and senescence and found that SNPs significantly associated with DTF for the E-haplotype were located at different positions from those identified with the whole dataset, and the most significantly associated SNP on Chr. 1 was found closest to AT1G33440, a major facilitator superfamily protein (Fig. 3 and *SI Appendix*, Fig. S9*B* and Tables S2, S7, and S8).

Taken together, our results suggest that correlations between seed chilling responses and reproductive timing may result in winter and summer annual life histories. *DOG1* variation is critical for cold induction of secondary dormancy that determines seasonal germination timing, and may also have pleiotropic effects on flowering time. However, a number of other loci are important to flowering-time variation, suggesting that the multivariate life-history syndromes we observe may be shaped by correlated selection at multiple loci in addition to *DOG1*.

***Arabidopsis* Germination Niche Today Is Explained by Current and Past Climate.** To understand the evolution of *DOG1*, we constructed a gene tree of the first exon of *DOG1*, which includes the self-binding domain (Fig. 4*A*). The earliest *DOG1* diversification represented in extant individuals is that of the ancestral and widely distributed ECCY haplotype, associated with traits that promote a winter annual life history. The tree topology showed a deep divergence within the ancestral ECCY clade, especially for relict, African, and Swedish accessions and a few anciently diverged haplotypes in China (Fig. 4*A* and *SI Appendix*, Figs. S6 and S7, and Table S9).

The rare ECSY haplotype arose within the ECCY clade, followed by EFSY, during periods of cooler climate after the last interglacial period (Fig. 4*A* and *SI Appendix*, Table S9); both of these derived haplotypes are geographically restricted to regions in Iberia with hot, dry summers and are associated with relict and Spanish accessions (Fig. 4 and *SI Appendix*, Figs. S6, S7, and S15). ECSY and EFSY likely originated in the Iberian refugium (46, 79–81), and may remain endemic in that region because they require hot, dry Mediterranean summers to persist.

D-SY and D-RY haplotypes are sister to one another, arising from an ECCY ancestor (Fig. 4*A*). Pairwise diversity within all D-clade individuals suggests this split occurred about 365 Kya (ranging from 281 to 450 Kya using differing assumptions about generation time) (*Materials and Methods*) during a period of climate cooling following an interglacial period (Fig. 4*B* and *SI Appendix*, Table S9), possibly during contraction into different refugia. However, we do not observe further diversification within the extant D-haplotypes until much later, potentially due to cycles of expansion and contraction out of refugia followed by local extinction. The extant, predominately Asian and east Eurasia D-SY clade diversified 115 Kya (88 to 141 Kya with varying generation times) during a cool period of the last interglacial period (Fig. 4*B* and *SI Appendix*, Table S9). Pairwise diversity among extant D-RY haplotypes suggests this west Eurasian and North American clade expanded recently 16 Kya (ranging 12 to 19 Kya), during rapid warming after the last glacial maximum (Fig. 4*B* and *SI Appendix*, Table S9), but the gene tree suggests an earlier origin (Fig. 4*A*). The extant distribution of D-RY in Iberia and Western Europe (*SI Appendix*, Figs. S6 and S7) suggests that this haplotype may have arisen and spread from the Iberian refugium.

These observations suggest that changing Pleistocene climate may have favored the rise and spread of the D-haplotypes, possibly through selection for associated traits promoting spring annual or rapid cycling life histories. The loss of cold-induced secondary dormancy in D-haplotypes, correlated with shorter life cycles, may have provided life-history flexibility facilitating persistence in changing climates and the invasion of new habitats. The ability to germinate in spring is advantageous in montane regions (17), such as Central Asia, and rapid cycling life histories may be favored in disturbed habitats with ample summer rainfall. Both D-SY and D-RY currently occupy climates with wetter summers than the E-haplotypes (*SI Appendix*, Fig. S15*B*). The wide Eurasian distribution of D-SY haplotypes suggests that their expanded germination niche may have facilitated postglacial expansion and spread of "nonrelict" genotypes out of central Asia into Europe and China (47, 82). The recent expansion of the D-RY clade is consistent with invasion of new ruderal habitats made available by the spread of agriculture through Europe (43, 82, 83). Our results show that derived loss-of-function D-haplotypes are more common in midlatitudes and England, and could have resulted from the nonrelict east–west expansion following human-mediated habitat disturbance. Relict populations in the Iberian Peninsula and Scandinavia might have preserved the ancestral *DOG1* E-haplotype that promotes a winter annual cycle, advantageous in more undisturbed habitats (81, 82). Moreover, the D-RY haplotype of *DOG1* is enriched in the invaded range of North America (*SI Appendix*, Fig. S6), where *A. thaliana* has arrived as a human commensal in the last 300 y (84). Although this could be due to a founder effect, phenological traits associated to this haplotype might also have played a role in invading disturbed habitats suitable for rapid cycling.

We caution that selection, geographic sampling, and demography could affect these estimates of age. But, with the exception of a more recent D-RY origin, these clade ages match the relative branching pattern on the gene tree. The origin of *DOG1* haplotype variants suggests that the germination niche of *A. thaliana* has been shaped by climate cycling throughout the Pleistocene, as the species range repeatedly expanded out of and contracted into glacial refugia (46, 79, 80) and the changing climate and physical environment filtered which individuals were able to persist (85). Novel germination strategies, measured via

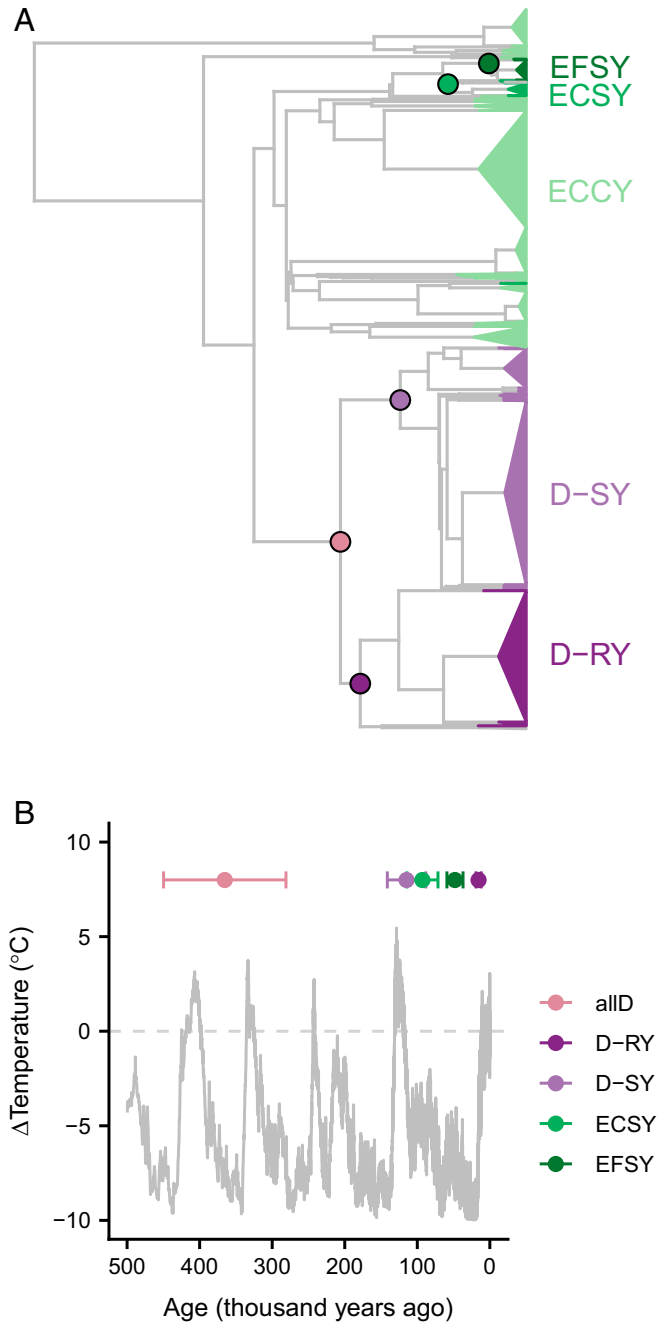

Fig. 4. (*A*) Bayesian phylogeny of *DOG1* haplotypes. Haplotype names represent amino acid sequence at the self-binding domain of *DOG1*. Triangle clade heights represent the number of individuals carrying a redundant sequence. The *A. lyrata* outgroup used to root the phylogeny is not shown. (*B*) Ages estimated from pairwise nucleotide divergence of *DOG1* haplotype groups with Pleistocene deuterium isotope records (103). Confidence intervals represent a range of generation times from 1.3 (±0.3) y per generation. Pairwise divergence of D-RY appears older in *B* as branch lengths are not weighted by number of individuals during tree-building (*Materials and Methods*). In both plots, points reflect divergence of haplotypes or groups of haplotypes. Temperatures are presented as deviations from the last 1,000-y average.

the origin of *DOG1* haplotypes, arose at key timepoints in the changing climate of the Quaternary. Thus, life-history strategies associated with *DOG1* haplotypes may in fact shape the population structure we see today on the landscape.

SNPs Associated with Chilling Responses and *DOG1* Functional Haplotypes Show a Signature of Climate Adaptation. To understand the climatic associations between genetic variants associated to various phenotypes, we used Gradient Forest, a tree-based machine-learning regression approach, to describe nonlinear turnover functions of allele frequencies along environmental gradients (40, 86, 87). We performed Gradient Forest regression analyses between environmental gradients and LD-pruned index SNPs from the 1,000 most highly associated SNPs from our GWA results for $PC1_{germ}$ (23 SNPs), $PC2_{germ}$ (9 SNPs), and the *DOG1* haplotypes. After accounting for spatial autocorrelation by including Moran's eigenvector map (MEM) variables, mean temperature of the wettest quarter (Bio8) explained the highest amount of turnover in allele frequency of SNPs associated to $PC2_{germ}$ and was among the top 3 predictors for $PC1_{germ}$ as well (Fig. 5 and *SI Appendix*, Table S10). Additionally, altitude, isothermality (Bio3), and temperature mean diurnal range (Bio2) predicted allele frequencies of both germination PCs (*SI Appendix*, Fig. S16 and Table S10). These environmental gradients structure turnover in index SNPs more than a set of 150 random SNPs (Fig. 5 and *SI Appendix*, Fig. S16), evidence of local adaptation to climate at these loci. Given that the E- and D-haplotypes are very old, our data suggest that ecotypes carrying both of these haplotypes could have arrived at locations throughout the range, but that environmental filtering and selection likely give rise to the geographic patterns that we see today. As a predominately selfing annual plant, *A. thaliana* populations are structured across the landscape, giving rise to geographic differences in ecologically relevant alleles (88); however, our results support local adaptation to climate, as allele turnover from index SNPs is more strongly structured by temperature and precipitation gradients than alleles from random SNPs.

The response-to-cold $PC2_{germ}$ cumulative importance function for mean temperature of the wettest quarter (Bio8) showed a threshold turnover in allele frequencies around 14 to 15 °C (Fig. 5), and alleles from SNPs associated to *DOG1* were among the ones with the highest importance values for this function (*SI Appendix*, Table S11). Accessions with the reference allele, with high dormancy release under chilling and no secondary dormancy, are found in these southern regions, while accessions with the alternative *DOG1* allele have strong secondary dormancy. This temperature threshold divides the winter growing season in southern regions (<14 to 15 °C) from a warm wet summer (>14 to 15 °C) growing season in northern latitudes, thus the shape of the allele turnover function suggests that germination traits might have been selected by these climatic gradients. Environmental predictors' importance varied at explaining the distribution of different *DOG1* haplotypes (*SI Appendix*, Fig. S17 and Tables S12 and S13). Allele distributions of *DOG1* might have resulted from its effect on germination responses to chilling and on other life-history traits as well. For example, 14 °C acts both as a temperature threshold for the induction of strong maternal effects on seed dormancy (11, 22) and for vernalization disruption in flowering time (89).

Conclusions

The seasonal germination niche shapes phenology and life history (1, 2, 5–8), and may be an essential component of adaptation to climate (11–13). Natural variation in dormancy and germination responses to seasonal environmental cues is often observed (21), and the genetic basis of variation in primary dormancy and after-ripening has been well studied in *A. thaliana*. However, much less is known about the genetic basis of natural variation in seed responses to seasonal chilling. Our data reveal a new axis of natural variation in germination and dormancy responses to cold that may drive the expression of winter annual vs. spring annual life histories. Moreover, this variation in response

to seasonal chilling was correlated with natural variation in flowering time, senescence, and after-ripening to form a range of life-history syndromes. GWA identified several loci associated with natural variation in chilling responses, including a known functional variant of the candidate gene *DOG1*, which showed a geographic signature of adaptation to present-day climate. A phylogeny of *DOG1* haplotypes revealed ancient divergence of these functional variants associated with past periods of climate change. Thus, natural variation in the *A. thaliana* germination niche is shaped by both past and present climate.

Materials and Methods

Seed Bulking. Seeds from 559 *A. thaliana* fully sequenced accessions from the 1001 Genomes Project (ABRC stocks) were stratified in 0.15% agar at 4 °C for 7 d. Seeds were sown in soil and allowed to germinate for 10 d; seedlings were vernalized for 6 wk at 4 °C with a 12/12-h photoperiod and were grown at 14 °C with a 12/12-h photoperiod in a walk-in growth chamber (Conviron E7/2 Controlled Environment Chamber). Two individuals of each accession were grown in the same chamber (a total of 1,118 plants), planted 2 wk apart in 2 temporal replicate blocks. Plants were watered twice a week with nutrient water until they showed 50 to 60% ripe siliques and seeds were individually harvested from each plant when 70 to 80% of the siliques were ripe. Fresh harvested seeds from each plant were immediately used for 2 different experiments: 1) Cold stratification germination experiments and 2) dry seed storage germination experiments to test for seed after-ripening times. Phenology variables were recorded on the maternal plants including DTF and DTS. DTF varied from 70 to 180 d while DTS ranged from 151 to 448 d (Fig. 2) (90).

Control and Dark Cold Stratification Germination Experiments. Fresh seeds were stored in a drying box with desiccant in the laboratory for 11 d, then assigned to 13 experimental treatments. Base germination of imbibed fresh seeds (control) was assessed under germination-inducing conditions after 7 d at 22 °C and 12/12-h photoperiod (80-μmols m s^{-1} light). We also exposed seeds to 12 stratification treatments. Seeds for these treatments were imbibed and stratified in boxes in dark chambers for 4, 8, 11, 15, 22, and 32 d at 4 °C (simulating winter temperature cues) or 10 °C (simulating early spring and autumn temperature cues) (11, 26). After these different cold treatments, dark-imbibed stratified seeds were put under the same germination-inducing conditions as the control seeds. Total germination was manually scored from photographs (i.e., radicle protrusion) after 7 d. Each experimental replicate consisted of 25 to 30 fresh seeds scattered on a Petri plate (60 × 15 mm) with blue germination paper (Blue Seed Germination Blotter, Anchor Paper) and 35 mL of a 2% PPM solution (Plant Preservative Mixture, Caisson Laboratories). PPM was added to prevent fungal growth during the 4- to 32-d stratification period (91). There were 2 replicates for each accession × treatment combination, for a total of 14,534 Petri plates scored. Germination experiments were carried out continuously for over a year, as fresh seed from individual plants were harvested when seeds matured across over 500 d, and not during a single month period (90).

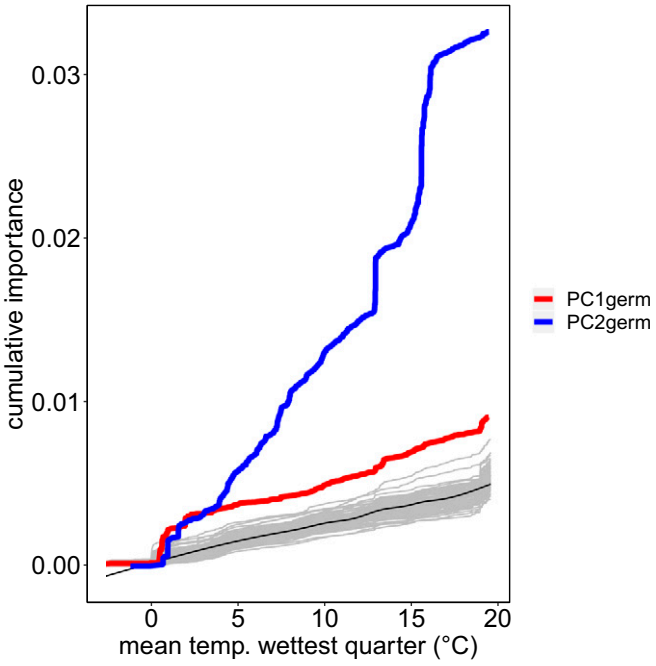

Fig. 5. Average cumulative importance functions of mean temperature of the wettest quarter for the index SNPs of the top 100 SNPs with the highest association with the first 2 PCs of germination for 752 accessions. Genomic controls are shown in grey, representing the 100 cumulative functions of 150 random SNPs with MAF > 0.01.

Dry Seed Storage Germination Experiments. Additional germination experiments were done to test for seed after-ripening. Fresh seeds were stored under dry laboratory conditions and were tested for total germination every 6 wk until 75% germination was recorded in 2 consecutive tests, which we considered as after-ripened seeds. Some accessions were after-ripened after 6 wk while some others showed little evidence of after-ripening even after dry storage in the laboratory for up to 788 d. Germination induction conditions were the same as for the cold stratification experiment. Seed after-ripening was assessed by calculating the number of days to 50% germination by fitting a polynomial regression to the dry-stored seed germination data for each ecotype ($DSDS_{50}$) (92).

Data Analyses.

Germination strategies (PC_{germ}). We excluded accessions that had misleading location data due to being misidentified according to ref. 93. We ran a PCA on mean percent germination from the 2 plantings for the base and cold treatment logit-transformed germination data (prcomp function in R). PCA scores and loadings were rescaled to the axes SD.

Germination responses and phenological associations. We tested for correlations between the first 2 PCs of germination variables, after-ripening ($DSDS_{50}$), and phenology variables of the mother plants (i.e., DTF and DTS; Pearson's product-moment correlation, cor.test function in R).

GWA on germination and phenology traits. GWA analyses were performed on the scores of the first 2 germination PCs, base germination, $DSDS_{50}$, DTF, and DTS to identify associated genetic polymorphisms using a univariate linear mixed model (LMM) corrected for population structure, implemented in GEMMA (94). Genotypes from sampled individuals ($n = 498$) were obtained from the 1001 Genomes Project (3,483,598 SNPs, minor allele frequency [MAF] 0.01, imputed genotypes) (43). Permutation-based thresholds were calculated for each trait by running a GWA 100 times with phenotypes permuted over genotypes and getting the average *P* value of the top fifth quantile from each analysis. We annotated SNPs to the closest TAIR10 gene (distanceToNearest function in GenomicRanges) with SNPs >1-kb distance from a gene annotated as intergenic regions. We also used TAGGIT to annotate the 1,000 most highly associated SNPs with respect to seed-specific gene ontology categories (95). GWA results were used to group SNPs into LD blocks. SNPs that were associated at a significance threshold of $P \leq 0.0001$ and that were not included in other LD blocks were selected as index SNPs. LD blocks around the index SNPs were defined by all other associated SNPs at a significant threshold of $P \leq 0.01$ that were in LD with the index SNP ($r^2 = 0.50$) within a physical distance of 250 kb (default settings from clump command, PLINK).

DOG1 functional haplotypes dated phylogeny. A 3-bp indel known to affect DOG1 function (53) is not genotyped in the available 1001 Genomes Project vcf files. In order to characterize our accessions at this known functional site, we mapped whole-genome resequencing reads from 1,135 *A. thaliana* accessions from the 1001 Genomes Consortium (43), 64 African *A. thaliana* accessions (45), and 118 Chinese *A. thaliana* accessions (46) to the primary cDNA of *DOG1* (AT5G45830.1) using bwa-mem v0.7.12-r1039 (96). We retained mapped reads and their pair, and assembled these using phrap v1.090518 (http://www.phrap.org). We aligned assemblies to exons of *DOG1* using MAFFT v7.27 (97), and extracted regions from the haplotype at amino acid positions 13 to 16, as defined by ref. 53. We filtered minor alleles with fewer than 2 individuals, which eliminated 3 rare haplotypes: C-CY, D-CY, and EYSY. Due to alignment and assembly issues, we recovered sequence for 972 individuals. *DOG1* local reassemblies of the 393 bp of exon 1 were pruned for redundant sequences, resulting in 74 unique individual sequences (98, 99) used to generate a Bayesian likelihood phylogeny in BEAST 2.5.2 with a strict clock and constant-sized coalescent prior, using an *Arabidopsis lyrata* and *Capsella rubella DOG1* sequence as outgroups. We ran BEAST for 10 million generations, confirmed Markov chain Monte Carlo convergence by eye, discarded 10% of trees as burn-in, and summarized trees with a maximum clade credibility tree. Attempts to run BEAST including

all 972 tips, many of which were redundant sequences, failed to converge. Admixture group assignments from the 1001 Genomes Consortium (43) were used to display the geographic distribution of the haplotypes on the tree, with "Africa" used for samples sourced from Durvasula et al. (45) and "China" used for samples sourced from Zou et al. (46). To estimate the timing of origin of each *DOG1* haplotype, we calculated average pairwise genetic diversity within the exon 1 sequence of each haplotype, using pegas (100), and converted to years using a mutation rate of 7.1e-9 mutations per generation (101) ($T = \pi/2u$, where T is time in generations, π is average pairwise diversity, and u is the per-generation mutation rate). As differences in life history can generate variation in generation times, we must make assumptions about generation time to directly relate generations to absolute time. Although *A. thaliana* is an annual plant, models suggest all environments except the southernmost Spanish population have generation times longer than 1 y (up to 4 y per generation) (7), and seed bank dynamics in northern Scandinavian populations estimate 1.3 y per generation (102). These differences in life history could be directly impacted by germination behavior. We use the experimental value of 1.3 (±0.3) years per generation (102) for all samples to estimate absolute time, but caution that relaxing this assumption can give rise to differing relationships to climatic events. Similarly, differences between the mutation rate and substitution rate (36) would affect the relationship to absolute climatic events. We related these dates to deuterium records from Jouzel et al. (103) in Fig. 4*B*. Scripts and files are available at https://github.com/mcstitzer/martinez-berdeja_dog1.

***DOG1* functional haplotypes associations with germination and phenology.** We used an LMM with population structure as a covariate (relmatLmer function in lme4qtl package in R) to analyze $PC1_{germ}$, $PC2_{germ}$, DTF, and DTS differences among of *DOG1* haplotypes ($n = 387$). Additionally, we analyzed regional phenotypic differences among *DOG1* haplotypes from Spanish ($n = 75$) and Swedish ($n = 125$) accessions.

Genomic environmental associations. We ran Gradient Forest (gradientForest R package) to find the environmental variables that best explained spatial variation in SNP alleles identified by GWA. We selected candidate SNPs to run Gradient Forest by LD pruning the 100 most highly associated SNPs from each GWA output (40, 86). We included genotypes from $n = 752$ accessions excluding those outside longitude −15 to 90 and randomly including 1 accession from the ones with identical location data. We used altitude, Bioclim 1 to 19 climate variables from the location of origin of each accession (WorldClim) (104), and included 23 MEM variables to account for spatial autocorrelation (40). Determining the number of MEMs necessary to account for population structure is not clear. We used a similar number of MEMs as the environmental variables used as environmental predictors, which corresponded to 5% of the MEMs calculated using adespatial R package following Jansen et al. (87). The top-ranked environmental variables were robust to inclusion of different numbers of MEMs. To assess whether the observed environmental associations were more extreme than those expected due to neutral processes, we tested a set of a random SNPs, resampling 100 times. These genomic controls were obtained by running Gradient Forest on 150 random SNPs, with a MAF > 0.01 and a similar LD ($r^2 = 0.5$ in a 250 k window that initiated every 10 SNPs using Plink). We also ran Gradient Forest on the different *DOG1* haplotypes to find the environmental variables that best explained haplotype distribution along environmental gradients (GradientForest R package). For this analysis we used the accessions from 1001 Genomes for which we assembled *DOG1* haplotypes and selected a set of $n = 516$ accessions, excluding the ones located outside longitude −15 to 90 and including a single representative from repeated location data.

Data availability. Germination, phenology and *DOG1* haplotype data are available on Dryad repository and AraPheno. *DOG1* haplotype data and code are also available on GitHub (https://github.com/mcstitzer/martinez-berdeja_dog1).

ACKNOWLEDGMENTS. We thank Kent Bradford, Moises Exposito-Alonso, Arthur Korte, Jeff Ross-Ibarra, Liana Burghardt, Emily Josephs, and Megan Bontrager for valuable advice and comments on the manuscript; and Mireille Caton-Darby, Holly Addington, Danielle Ethington, Felicia Wong, Josh Leung, Bryan González, Helena Bayat, Lydia Eldridge, and Jasneek Attwal for growing the accessions and carrying out the germination experiments. This work was supported by a Conacyt Postdoctoral fellowship (to A.M.-B.); National Science Foundation Grants DEB-1447203 and DEB-1754102 (to J.S.); and the University of California, Davis.

1. K. Donohue, Germination timing influences natural selection on life-history characters in *Arabidopsis thaliana*. *Ecology* **83**, 1006–1016 (2002).
2. K. Donohue, Niche construction through phenological plasticity: Life history dynamics and ecological consequences. *New Phytol.* **166**, 83–92 (2005).
3. L. F. Galloway, Parental environmental effects on life history in the herbaceous plant *Campanula americana*. *Ecology* **82**, 2781–2789 (2001).
4. A. M. Wilczek *et al.*, Effects of genetic perturbation on seasonal life history plasticity. *Science* **323**, 930–934 (2009).
5. X. Huang *et al.*, The earliest stages of adaptation in an experimental plant population: Strong selection on QTLS for seed dormancy. *Mol. Ecol.* **19**, 1335–1351 (2010).
6. G. C. Chiang *et al.*, Pleiotropy in the wild: The dormancy gene *DOG1* exerts cascading control on life cycles. *Evolution* **67**, 883–893 (2013).
7. L. T. Burghardt, C. J. Metcalf, A. M. Wilczek, J. Schmitt, K. Donohue, Modeling the influence of genetic and environmental variation on the expression of plant life cycles across landscapes. *Am. Nat.* **185**, 212–227 (2015).
8. L. T. Burghardt, B. R. Edwards, K. Donohue, Multiple paths to similar germination behavior in *Arabidopsis thaliana*. *New Phytol.* **209**, 1301–1312 (2016).
9. J. R. Gremer, C. J. Wilcox, A. Chiono, E. Suglia, J. Schmitt, Germination timing and chilling exposure create contingency in life history and influence fitness in the native wildflower *Streptanthus tortuosus*. *J. Ecol.* **108**, 239–255 (2020).
10. L. F. Galloway, J. R. Etterson, Transgenerational plasticity is adaptive in the wild. *Science* **318**, 1134–1136 (2007).
11. A. Montesinos-Navarro, F. X. Picó, S. J. Tonsor, Clinal variation in seed traits influencing life cycle timing in *Arabidopsis thaliana*. *Evolution* **66**, 3417–3431 (2012).
12. D. S. Vidigal *et al.*, Altitudinal and climatic associations of seed dormancy and flowering traits evidence adaptation of annual life cycle timing in *Arabidopsis thaliana*. *Plant Cell Environ.* **39**, 1737–1748 (2016).
13. F. M. Postma, J. Ågren, Early life stages contribute strongly to local adaptation in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 7590–7595 (2016).
14. S. Kalisz, Variable selection on the timing of germination in *Collinsia verna* (Scrophulariaceae). *Evolution* **40**, 479–491 (1986).
15. K. Donohue *et al.*, The evolutionary ecology of seed germination of *Arabidopsis thaliana*: Variable natural selection on germination timing. *Evolution* **59**, 758–770 (2005).
16. T. M. Korves *et al.*, Fitness effects associated with the major flowering time gene *FRIGIDA* in *Arabidopsis thaliana* in the field. *Am. Nat.* **169**, E141–E157 (2007).
17. F. X. Picó, Demographic fate of *Arabidopsis thaliana* cohorts of autumn- and spring-germinated plants along an altitudinal gradient. *J. Ecol.* **100**, 1009–1018 (2012).
18. L. T. Burghardt, C. J. Metcalf, K. Donohue, A cline in seed dormancy helps conserve the environment experienced during reproduction across the range of *Arabidopsis thaliana*. *Am. J. Bot.* **103**, 47–59 (2016).
19. M. Debieu *et al.*, Co-variation between seed dormancy, growth rate and flowering time changes with latitude in *Arabidopsis thaliana*. *PLoS One* **8**, e61075 (2013).
20. A. Marcer *et al.*, Temperature fine-tunes Mediterranean *Arabidopsis* thaliana life-cycle phenology geographically. *Plant Biol (Stuttg)* **20** (suppl. 1), 148–156 (2018).
21. C. C. Baskin, J. M. Baskin, *Seeds. Ecology, Biogeography, and Evolution of Dormancy and Germination* (San Diego Academic Press, ed. 2, 2014).
22. S. Penfield, V. Springthorpe, Understanding chilling responses in *Arabidopsis* seeds and their contribution to life history. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **367**, 291–297 (2012).
23. S. Footitt, Z. Huang, H. A. Clay, A. Mead, W. E. Finch-Savage, Temperature, light and nitrate sensing coordinate *Arabidopsis* seed dormancy cycling, resulting in winter and summer annual phenotypes. *Plant J.* **74**, 1003–1015 (2013).
24. W. E. Finch-Savage, S. Footitt, Seed dormancy cycling and the regulation of dormancy mechanisms to time germination in variable field environments. *J. Exp. Bot.* **68**, 843–856 (2017).
25. J. R. Gremer, S. Kimball, D. L. Venable, Within-and among-year germination in Sonoran Desert winter annuals: Bet hedging and predictive germination in a variable environment. *Ecol. Lett.* **19**, 1209–1218 (2016).
26. G. A. Auge *et al.*, Secondary dormancy dynamics depends on primary dormancy status in *Arabidopsis thaliana*. *Seed Sci. Res.* **25**, 230–246 (2015).
27. L. Bentsink, J. Jowett, C. J. Hanhart, M. Koornneef, Cloning of *DOG1*, a quantitative trait locus controlling seed dormancy in *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 17042–17047 (2006).
28. L. Bentsink, M. Koornneef, Seed dormancy and germination. *Arabidopsis Book* **6**, e0119 (2008).
29. S. Kendall, S. Penfield, Maternal and zygotic temperature signalling in the control of seed dormancy and germination. *Seed Sci. Res.* **22**, S23–S29 (2012).
30. M. Chen *et al.*, Maternal temperature history activates Flowering Locus T in fruits to control progeny dormancy according to time of year. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 18787–18792 (2014).
31. H. He *et al.*, Interaction between parental environment and genotype affects plant and seed performance in *Arabidopsis*. *J. Exp. Bot.* **65**, 6603–6615 (2014).
32. S. Footitt, I. Douterelo-Soler, H. Clay, W. E. Finch-Savage, Dormancy cycling in *Arabidopsis* seeds is controlled by seasonally distinct hormone-signaling pathways. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 20236–20241 (2011).
33. S. Footitt, H. A. Clay, K. Dent, W. E. Finch-Savage, Environment sensing in spring-dispersed seeds of a winter annual *Arabidopsis* influences the regulation of dormancy to align germination potential with seasonal changes. *New Phytol.* **202**, 929–939 (2014).
34. W. E. Finch-Savage, G. Leubner-Metzger, Seed dormancy and the control of germination. *New Phytol.* **171**, 501–523 (2006).
35. J. D. Bewley, K. Bradford, H. Hilhorst, H. Nonogaki, *Seeds: Physiology of Development, Germination and Dormancy* (Springer, ed. 3, 2013).

36. M. Exposito-Alonso *et al*., The rate and potential relevance of new mutations in a colonizing plant lineage. *PLoS Genet.* **14**, e1007155 (2018).
37. M. Takou *et al*., Linking genes with ecological strategies in *Arabidopsis thaliana*. *J. Exp. Bot.* **70**, 1141–1151 (2019).
38. J. R. Etterson, R. G. Shaw, Constraint to adaptive evolution in response to global warming. *Science* **294**, 151–154 (2001).
39. A. Fournier-Level *et al*., A map of local adaptation in *Arabidopsis thaliana*. *Science* **334**, 86–89 (2011).
40. M. C. Fitzpatrick, S. R. Keller, Ecological genomics meets community-level modelling of biodiversity: Mapping the genomic landscape of current and future environmental adaptation. *Ecol. Lett.* **18**, 1–16 (2015).
41. M. Exposito-Alonso *et al*., Genomic basis and evolutionary potential for extreme drought adaptation in *Arabidopsis thaliana*. *Nat. Ecol. Evol.* **2**, 352–358 (2018).
42. S. Atwell *et al*., Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. *Nature* **465**, 627–631 (2010).
43. 1001 Genomes Consortium, 1,135 Genomes reveal the global pattern of polymorphism in *Arabidopsis thaliana*. *Cell* **166**, 481–491 (2016).
44. J. B. Beck, H. Schmuths, B. A. Schaal, Native range genetic variation in *Arabidopsis thaliana* is strongly geographically structured and reflects Pleistocene glacial dynamics. *Mol. Ecol.* **17**, 902–915 (2008).
45. A. Durvasula *et al*., African genomes illuminate the early history and transition to selfing in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 5213–5218 (2017).
46. Y. P. Zou *et al*., Adaptation of *Arabidopsis thaliana* to the Yangtze River basin. *Genome Biol.* **18**, 239 (2017).
47. C. W. Hsu, C. Y. Lo, C. R. Lee, On the postglacial spread of human commensal *Arabidopsis thaliana*: Journey to the East. *New Phytol.* **222**, 1447–1457 (2019).
48. J. R. Stinchcombe *et al*., A latitudinal cline in flowering time in *Arabidopsis thaliana* modulated by the flowering time gene *FRIGIDA*. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 4712–4717 (2004).
49. F. Bouché, G. Lobet, P. Tocquin, C. Périlleux, FLOR-ID: An interactive database of flowering-time gene networks in *Arabidopsis thaliana*. *Nucleic Acids Res.* **44**, D1167–D1171 (2016).
50. E. Kerdaffrec *et al*., Multiple alleles at a single locus control seed dormancy in Swedish *Arabidopsis*. *eLife* **5**, e22502 (2016).
51. I. Kronholm, F. X. Picó, C. Alonso-Blanco, J. Goudet, J. de Meaux, Genetic basis of adaptation in *Arabidopsis thaliana*: Local adaptation at the seed dormancy QTL *DOG1*. *Evolution* **66**, 2287–2302 (2012).
52. F. M. Postma, J. Ågren, Maternal environment affects the genetic basis of seed dormancy in *Arabidopsis thaliana*. *Mol. Ecol.* **24**, 785–797 (2015).
53. K. Nakabayashi, M. Bartsch, J. Ding, W. J. Soppe, Seed dormancy in *Arabidopsis* requires self-binding ability of *DOG1* protein and the presence of multiple isoforms generated by alternative splicing. *PLoS Genet.* **11**, e1005737 (2015).
54. D. Tabas-Madrid *et al*., Genome-wide signatures of flowering adaptation to climate temperature: Regional analyses in a highly diverse native range of *Arabidopsis thaliana*. *Plant Cell Environ.* **41**, 1806–1820 (2018).
55. B. R. Ni, K. J. Bradford, Germination and dormancy of abscisic acid- and gibberellin-deficient mutant tomato (*Lycopersicon esculentum*) seeds. *Plant Physiol.* **101**, 607–617 (1993).
56. G. Leubner-Metzger, Functions and regulation of β-1,3-glucanases during seed germination, dormancy release and after-ripening. *Seed Sci. Res.* **13**, 17–34 (2003).
57. A. Voegele, A. Linkies, K. Müller, G. Leubner-Metzger, Members of the gibberellin receptor gene family *GID1* (*GIBBERELLIN INSENSITIVE DWARF1*) play distinct roles during *Lepidium sativum* and *Arabidopsis thaliana* seed germination. *J. Exp. Bot.* **62**, 5131–5147 (2011).
58. G. Née *et al*., *DELAY OF GERMINATION1* requires PP2C phosphatases of the ABA signalling pathway to control seed dormancy. *Nat. Commun.* **8**, 72 (2017).
59. K. Nakabayashi *et al*., The time required for dormancy release in *Arabidopsis* is determined by *DELAY OF GERMINATION1* protein levels in freshly harvested seeds. *Plant Cell* **24**, 2826–2838 (2012).
60. H. Huo, S. Wei, K. J. Bradford, *DELAY OF GERMINATION1* (*DOG1*) regulates both seed dormancy and flowering time through microRNA pathways. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E2199–E2206 (2016).
61. A. Linkies *et al*., Ethylene interacts with abscisic acid to regulate endosperm rupture during germination: A comparative approach using *Lepidium sativum* and *Arabidopsis thaliana*. *Plant Cell* **21**, 3803–3822 (2009).
62. K. Graeber *et al*., *DELAY OF GERMINATION 1* mediates a conserved coat-dormancy mechanism for the temperature- and gibberellin-dependent control of seed germination. *Proc. Natl. Acad. Sci. U.S.A.* **111**, E3571–E3580 (2014).
63. G. C. Chiang *et al*., *DOG1* expression is predicted by the seed-maturation environment and contributes to geographical variation in germination in *Arabidopsis thaliana*. *Mol. Ecol.* **20**, 3336–3349 (2011).
64. Y. Li, Y. Huang, J. Bergelson, M. Nordborg, J. O. Borevitz, Association mapping of local climate-sensitive quantitative trait loci in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 21199–21204 (2010).
65. M. A. Taylor *et al*., Interacting effects of genetic variation for seed dormancy and flowering time on phenology, life history, and fitness of experimental *Arabidopsis thaliana* populations over multiple generations in the field. *New Phytol.* **216**, 291–302 (2017).
66. J. R. Gremer, D. L. Venable, Bet hedging in desert winter annual plants: Optimal germination strategies in a variable environment. *Ecol. Lett.* **17**, 380–387 (2014).
67. N. Price *et al*., Combining population genomics and fitness QTLs to identify the genetics of local adaptation in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 5028–5033 (2018).
68. J. F. Golz *et al*., Layers of regulation—Insights into the role of transcription factors controlling mucilage production in the *Arabidopsis* seed coat. *Plant Sci.* **272**, 179–192 (2018).
69. H. Nonogaki, O. H. Gee, K. J. Bradford, A germination-specific endo-β-mannanase gene is expressed in the micropylar endosperm cap of tomato seeds. *Plant Physiol.* **123**, 1235–1246 (2000).
70. F. Chen, K. J. Bradford, Expression of an expansin is associated with endosperm weakening during tomato seed germination. *Plant Physiol.* **124**, 1265–1274 (2000).
71. W. Chen *et al*., Expression profile matrix of *Arabidopsis* transcription factor genes suggests their putative functions in response to environmental stresses. *Plant Cell* **14**, 559–574 (2002).
72. Y. Yamauchi *et al*., Activation of gibberellin biosynthesis and response pathways by low temperature during imbibition of *Arabidopsis thaliana* seeds. *Plant Cell* **16**, 367–378 (2004).
73. S. L. Kendall *et al*., Induction of dormancy in *Arabidopsis* summer annuals requires parallel regulation of *DOG1* and hormone metabolism by low temperature and CBF transcription factors. *Plant Cell* **23**, 2568–2580 (2011).
74. H. Fedak *et al*., Control of seed dormancy in *Arabidopsis* by a cis-acting noncoding antisense transcript. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E7846–E7855 (2016).
75. R. Yatusevich *et al*., Antisense transcription represses *Arabidopsis* seed dormancy QTL *DOG1* to regulate drought tolerance. *EMBO Rep.* **18**, 2186–2196 (2017).
76. G. A. Auge, S. Penfield, K. Donohue, Pleiotropy in developmental regulation by flowering-pathway genes: Is it an evolutionary constraint? *New Phytol.* **224**, 55–70 (2019).
77. M. Togninalli *et al*., The AraGWAS catalog: A curated and standardized *Arabidopsis thaliana* GWAS catalog. *Nucleic Acids Res.* **46**, D1150–D1156 (2018).
78. Y. Zan, Ö. Carlborg, A polygenic genetic architecture of flowering time in the worldwide *Arabidopsis thaliana* population. *Mol. Biol. Evol.* **36**, 141–154 (2019).
79. H. P. Comes, J. W. Kadereit, The effect of Quaternary climatic changes on plant distribution and evolution. *Trends Plant Sci.* **3**, 432–438 (1998).
80. T. F. Sharbel, B. Haubold, T. Mitchell-Olds, Genetic isolation by distance in *Arabidopsis thaliana*: Biogeography and postglacial colonization of Europe. *Mol. Ecol.* **9**, 2109–2118 (2000).
81. B. Toledo, A. Marcer, B. Méndez-Vigo, C. Alonso-Blanco, F. X. Picó, An ecological history of the relict genetic lineage of *Arabidopsis thaliana*. *Environ. Exp. Bot.*, 10.1016/j.envexpbot.2019.103800 (2019).
82. C. R. Lee *et al*., On the post-glacial spread of human commensal *Arabidopsis thaliana*. *Nat. Commun.* **8**, 14458 (2017).
83. O. François, M. G. Blum, M. Jakobsson, N. A. Rosenberg, Demographic history of european populations of *Arabidopsis thaliana*. *PLoS Genet.* **4**, e1000075 (2008).
84. J. Hagmann *et al*., Century-scale methylome stability in a recently diverged *Arabidopsis thaliana* lineage. *PLoS Genet.* **11**, e1004920 (2015).
85. G. M. Hewitt, Some genetic consequences of ice ages, and their role in divergence and speciation. *Biol. J. Linn. Soc. Lond.* **58**, 247–276 (1996).
86. R. A. Bay *et al*., Genomic signals of selection predict climate-driven population declines in a migratory bird. *Science* **359**, 83–86 (2018).
87. A. Jansen van Rensburg, M. Cortazar-Chinarro, A. Laurila, J. Van Buskirk, Adaptive genomic variation associated with environmental gradients along a latitudinal cline in Rana temporaria. *BioRxiv* https://doi.org/10.1101/427872 (27 September 2018).
88. C. Toomajian *et al*., A nonparametric test reveals selection for rapid flowering in the *Arabidopsis* genome. *PLoS Biol.* **4**, e137 (2006).
89. S. Duncan *et al*., Seasonal shift in timing of vernalization as an adaptation to extreme winter. *eLife* **4**, e06620 (2015).
90. A. Martínez-Berdeja, J. Schmitt, arabidopsis_germ_phen_chilling_DATA.xls. Dryad. https://doi.org/10.25338/B8VS4P. Deposited 10 December 2019.
91. H. Huo *et al*., Rapid identification of lettuce seed germination mutants by bulked segregant analysis and whole genome sequencing. *Plant J.* **88**, 345–360 (2016).
92. C. Alonso-Blanco, L. Bentsink, C. J. Hanhart, H. Blankestijn-de Vries, M. Koornneef, Analysis of natural allelic variation at seed dormancy loci of *Arabidopsis thaliana*. *Genetics* **164**, 711–729 (2003).
93. A. E. Anastasio *et al*., Source verification of mis-identified Arabidopsis thaliana accessions. *Plant J.* **67**, 554–566 (2011).
94. X. Zhou, P. Carbonetto, M. Stephens, Polygenic modeling with bayesian sparse linear mixed models. *PLoS Genet.* **9**, e1003264 (2013).
95. E. Carrera *et al*., Gene expression profiling reveals defined functions of the ATP-binding cassette transporter COMATOSE late in phase II of germination. *Plant Physiol.* **143**, 1669–1679 (2007).
96. H. Li, Aligning sequences reads, clone sequences and assembly contigs with BWA-MEM. arXiv:1303.3997 (26 May 2013).
97. K. Katoh, D. M. Standley, MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
98. M. C. Stitzer, dog1.may22.uniq.txt. Dryad. https://doi.org/10.25338/B8VS4P. Deposited 10 December 2019.
99. M. C. Stitzer, Scripts to generate DOG1 haplotypes and understand their evolution for Martinez-Berdeja et al., 2019. GitHub. https://github.com/mcstitzer/martinez-berdeja_dog1. Deposited 10 December 2019.
100. E. Paradis, pegas: An R package for population genetics with an integrated-modular approach. *Bioinformatics* **26**, 419–420 (2010).
101. S. Ossowski *et al*., The rate and molecular spectrum of spontaneous mutations in *Arabidopsis thaliana*. *Science* **327**, 92–94 (2010).
102. M. Falahati-Anbaran, S. Lundemo, H. K. Stenøien, Seed dispersal in time can counteract the effect of gene flow between natural populations of *Arabidopsis thaliana*. *New Phytol.* **202**, 1043–1054 (2014).
103. J. Jouzel *et al*., Orbital and millennial Antarctic climate variability over the past 800,000 years. *Science* **317**, 793–796 (2007).
104. R. J. Hijmans, S. E. Cameron, J. L. Parra, P. G. Jones, A. Jarvis, Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* **25**, 1965–1978 (2005).